UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/695,077	10/25/2000	PHILIP NEIL GARNER	1263.1881	1606

5514 7590 04/06/2005

FITZPATRICK CELLA HARPER & SCINTO
30 ROCKEFELLER PLAZA
NEW YORK, NY 10112

EXAMINER	
LERNER, MARTIN	

ART UNIT	PAPER NUMBER
2654	

DATE MAILED: 04/06/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/695,077	GARNER ET AL.	
	Examiner	Art Unit	
	Martin Lerner	2654	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *19 July 2004 and 06 December 2004*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1 to 95 and 97 to 102* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-5,20-25,30-39,46,48-52,67-73,77-86,90,93 and 97-102* is/are rejected.

7)☒ Claim(s) *6-19,26-29,40-45,47,53-66,74-76,87-89,91,92,94 and 95* is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☒ All b)☐ Some * c)☐ None of:

1.☒ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date *6/08/04 & 12/08/04*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Claim Rejections - 35 USC § 102

1. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

2. Claims 1 to 5, 21 to 25, 30 to 32, 37 to 39, 48 to 52, 68 to 73, 77 to 79, 84 to 86, 90, 97, 99, and 101 are rejected under 35 U.S.C. 102(b) as being anticipated by *Chou et al.*

Regarding independent claims 1, 48, 97, 99, and 101, *Chou et al.* discloses an apparatus, method, and instructions, comprising:

"a receiver operable to receive an input signal" – an input of an unknown speech string 18 (an utterance) of words is received from a microphone (column 4, lines 34 to 35: Figure 1);

"a recognition processor operable to compare said input signal with stored label models to generate a recognized sequence of labels in said input signal and confidence data representative of the confidence that the recognized sequence of labels is representative of said input signal" – recognition processor 10 receives the input, accesses the recognition database 12, scores the unknown speech string of words against the recognition models in the recognition database 12, and generates a

hypothesis string signal 20; verification processor 16 receives the hypothesis string, and generates a confidence measure signal 22 (column 4, lines 34 to 51: Figure 1);

"a similarity measure calculator operable to compare said recognized sequence of labels received from said recognition processor with a stored sequence of labels using a combination of i) predetermined confusion data which defines confusability between different labels, and ii) said confidence data received from the recognition processor and representative of the confidence that said received recognized sequence of labels is representative of the input signal, to provide a measure of the similarity between the recognized sequence of labels and the stored sequence of labels" – confidence score computation ("a similarity measure calculator") for a speech segment q relates a comparison between a word model score ("said confidence data") and scores computed with the anti-word model ("predetermined confusion data which defines confusability between different labels"); in Equation (2), $L(O_q;\Theta,l)$ is "the measure of similarity" calculated by the similarity measure calculator, $g_l(O_q) = log\ p(O_q|\theta_l^{(k)})$ is "the confidence data" for the keyword hypothesis $\{\theta_l^{(k)}\}$, and $G_l(O_q)$ is the "predetermined confusion data which defines confusability" for anti-keywords $\{\theta_l^{(a)}\}$ which handle confusibility among keywords (column 8, lines 33 to 55: Figure 2).

Regarding claims 2 and 49, *Chou et al.* discloses the confidence measure is generated based upon data stored in verification database 16 (column 4, lines 34 to 51: Figure 1).

Regarding claims 3 and 50, *Chou et al.* discloses a word-based confidence score 34 (column 8, lines 40 to 55: Figure 2); each word is a "label" in a string of words being recognized.

Regarding claims 4 and 51, *Chou et al.* discloses string models are generated in an "N" best list ("a list of alternatives") by N-best string model generator 46 (column 6, line 45 to column 7, line 53: Figure 2; column 9, line 54 to column 10, line 14).

Regarding claims 5 and 52, *Chou et al.* discloses Viterbi alignment ("an aligner") of the input string, O, against the model sets for each given word string in the N-best string list (column 7, lines 8 to 15); average word-based confidence score processor 36 ("a combiner") performs mathematical averaging for each word segment signal of the hypothesis string to generate an average word-based confidence score signal ("said similarity measure") (column 5, lines 53 to 67: Figure 2).

Regarding claims 21 and 68, *Chou et al.* discloses Viterbi alignment ("an aligner") of the input string, O, against the models sets for each given word string in the N-best string list (column 7, lines 8 to 15); Viterbi alignment is "a dynamic programming technique".

Regarding claims 22 to 25, 69 to 72, and 90, *Chou et al.* discloses Viterbi alignment ("an aligner") of the input string, O, against the model sets for each given word string in the N-best string list (column 7, lines 8 to 15); implicitly, Viterbi alignment determines "progressively a plurality of possible alignments", generates scores for each given word in the N-best list, determines "an optimum alignment", and "combines the scores" for each word in the word string.

Regarding claims 30 to 32 and 77 to 79, *Chou et al.* discloses the input string is speech (column 4, lines 34 to 35: Figure 1), which is a time sequential audio signal of words.

Regarding claims 37, 38, 84, and 85, *Chou et al.* discloses confidence score computation for a speech segment q relates a comparison between a word model score ("said confidence data") and scores computed with the anti-word model ("said confusion data"); in Equation (2), $L(O_q;\Theta,l)$ is "the measure of similarity" calculated by the similarity measure calculator, $g_l(O_q) = log\ p(O_q|\theta_l^{(k)})$ is "the confidence data" for the keyword hypothesis $\{\theta_l^{(k)}\}$, and $G_l(O_q)$ is "the confusion data" for anti-keywords $\{\theta_l^{(a)}\}$ which handle confusibility among keywords (column 8, lines 33 to 55: Figure 2).

Regarding claims 39 and 86, *Chou et al.* discloses an average confidence score based on upon the average of word-based confidence scores (column 5, lines 53 to 67); an average confidence score is a normalization from each of the word-based confidence scores.

Regarding claim 73, *Chou et al.* discloses each of the words ("labels") in the unknown speech string ("each of the labels in said recognized sequence of labels") is scored against recognition models ("stored sequences of labels") in the recognition database 12 (column 4, lines 34 to 51: Figure 1).

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

4. Claims 20, 35, 36, 46, 67, 82, 83, 93, 98, 100, and 102 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Chou et al.* in view of *Aref et al.*

Concerning claims 20 and 67, *Chou et al.* omits an aligner operable to identify deletions and insertions. However, *Aref et al.* teaches an analogous art speech recognition system for correcting misspelled words in a string of text. (Column 1, Lines 32 to 52) Specifically, *Aref et al.* discloses detecting recognition errors as models from insertion errors and deletion errors. (Column 3, Lines 36 to 60) It is suggested that there are advantages to speed the search process and reduce the size of the database by correcting misrecognized or misspelled words with the search technique of *Aref et al.* (Column 1, Lines 52 to 61) It would have been obvious to one having ordinary skill in the art to incorporate the insertion and deletion error technique of *Aref et al.* into the word-based confidence score method of *Chou et al.* for the purpose of correcting misrecognitions with a high speed search process and reduced database size.

Concerning claims 35, 36, 82, and 83, *Chou et al.* omits mis-typing probabilities and mis-spelling probabilities based upon sub-word units. However, *Aref et al.* teaches an analogous art speech recognition system for correcting misspelled words in a string of text. (Column 1, Lines 32 to 52) Specifically, *Aref et al.* discloses probabilities for letters being recognized incorrectly, where letters are sub-word units, to estimate a measure of similarity between two words. (Column 4, Lines 1 to 59) Recognition errors are based upon typing errors, e.g. "airnmail" is mistakenly inserted for the word

"airmail". (Column 3, Lines 36 to 53) It is suggested that there are advantages to speed the search process and reduce the size of the database by correcting misrecognized or misspelled words with the search technique of *Aref et al.* (Column 1, Lines 52 to 61) It would have been obvious to one having ordinary skill in the art to utilize the mis-typing and mis-spelling technique of sub-word units taught by *Aref et al.* into the word-based confidence score method of *Chou et al.* for the purpose of correcting misrecognitions with a high speed search process and reduced database size.

Concerning claims 46, 93, 98, 100, and 102, *Chou et al.* omits an application of speech recognition to querying a database and obtaining information from the database, although this is a well known application for speech recognition systems, generally. However, *Aref et al.* teaches an analogous art speech recognition system for searching a database for recognized text by querying keywords. (Column 2, Lines 41 to 50) It is suggested that there are advantages to speed the search process and reduce the size of the database by correcting misrecognized or misspelled words with the search technique of *Aref et al.* (Column 1, Lines 52 to 61) It would have been obvious to one having ordinary skill in the art to apply the word-based confidence score method of *Chou et al.* to a retrieval system from a database of automatically recognized text as taught by *Aref et al.* for the purpose of correcting misrecognitions with a high speed search process and reduced database size.

5. Claims 33, 34, 80, and 81 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Chou et al.* in view of *Wheatley et al.*

Chou et al. discloses recognizing speech with word-based confidence scores, where the labels are words, but omits recognizing sub-word units and phonemes. However, it is a well known art recognized alternative in speech recognition to recognize phonemes, which are sub-word units, rather than words. *Wheatley et al.* teaches a related apparatus and method for speech recognition, where speech is recognized with Hidden Markov Models representing phonetic units instead of words. (Column 7, Lines 14 to 37) It is suggested that there is an advantage of representing real world, unscripted conversations. (Column 2, Lines 28 to 39) It would have been obvious to utilize sub-word phonetic units for the speech recognition system of *Chou et al.* as suggested by *Wheatley et al.* for the purpose of better recognizing real world, unscripted conversations.

Allowable Subject Matter

6. Claims 6 to 19, 26 to 29, 40 to 45, 47, 53 to 66, 74 to 76, 87 to 89, 91 to 92, and 94 to 95 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

Response to Arguments

7. Applicants' arguments filed 19 July 2004 and 06 December 2004 have been fully considered but they are not persuasive.

Applicants argue that *Chou et al.* fails to disclose the invention because the reference is based upon recognizing individual words, while the invention is operable to compare a sequence of labels with a stored sequence of labels. Applicants say a label may be a word or a phoneme. Applicants point to *Chou et al.*'s Equation (2), which they contend represents an individual word, or keyword, and not a sequence of words. Applicants admit that *Chou et al.*'s Equation (1) sums a confidence measure signal for each word to obtain a generated confidence measure signal for the whole of the hypothesized string of words. (Remarks, Page 33, of Amendment filed 19 July 2004) This is not persuasive.

Chou et al. repeatedly states that the recognition processor operates on a string of words. The recognition processor receives as input an unknown speech string 1 (an utterance) of words. The recognition processor 10 accesses the recognition database 12 in response to the unknown speech string 18 input and scores the unknown speech string of words against the recognition models in the recognition database 12. (Column 4, Lines 33 to 51: Figure 1) Applicants say their labels can be either words or phonemes. *Chou et al.* discloses strings of words, so each word is equivalent to a label, and a string of words corresponds to a sequence of labels. Moreover, *Chou et al.* compares an unknown speech string of words to recognition models. A speech string

represents a string of words, so it must follow that the models must correspond to word string models. Thus, *Chou et al.* expressly discloses a recognition processor that does not simply operate on individual words, but on strings of words and word models, corresponding to sequences of labels.

Similarly, Column 7, Lines 8 to 40, of *Chou et al.* discloses generating N-best string models. Those skilled in the art know a paradigmatic input phrase for a word string is: "Let's recognize speech." A correct recognition of the phrase can return, "Let's recognize speech," but an incorrect recognition produces, "Let's wreck a nice beach." "Let's recognize speech" and "Let's wreck a nice beach" would be present in a set of N phrases in an N-best list of a set of hypothesis word strings. Both phrases are based upon strings of individually valid words, but an effective speech recognition procedure needs to discriminate between a correctly recognized string of words and an incorrectly recognized string of words.

As admitted by Applicants, *Chou et al.* performs this procedure by generating a confidence score ("similarity measure") between a keyword hypothesis ("confidence data") and its competing alternative anti-keyword hypothesis, which handles confusibility, and then combines the contribution of the word-based confidence scores of the word signal segments to generate the string-based confidence measure for a hypothesis string signal. (Column 8, Lines 2 to 10; Column 8, Lines 33 to 55) Applicants contend that combining the individual words scores to generate a string score as in *Chou et al.* is somehow distinct from their procedure of using a combination of confusion data and confidence data for a received sequence of labels, as claimed.

This position is traversed, as it is maintained the processes are equivalent. Even supposing *Chou et al.* scores the words individually, and then adds the scores of the individual words to generate the score of the word string -- as Applicants note is indeed disclosed by *Chou et al.* -- *Chou et al.* still meets the limitations of the claims. The language of the claims, as drafted, does not distinguish over the procedure disclosed by *Chou et al.* Although the claims are interpreted in light of the specification, limitations from the specification are not read into the claims. See *In re Van Geuns*, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993).

Furthermore, Applicants have not shown that their claims should be interpreted as comparing entire words strings to generate confusion data and confidence data, as they contend should be the basis for distinguishing their invention over *Chou et al.* Applicants should cite specific pages and line numbers to support their interpretation from the Specification. Technically, one skilled in the art would expect that, in order to compare a string of keyword and anti-keyword hypotheses with models, one would necessarily need to first compare each word in the string individually. As a result, unless Applicants can show a contrary teaching from their Specification, it is believed that the mechanics of any comparison would necessarily require some initial word-based scoring before a string-based scoring could be possible, so that the procedures are equivalent.

Moreover, one skilled in the art would know that words implicitly are composed of phonemes for an utterance of a string of words and of word models, so it is implicit that

keyword and anti-keyword scoring is comparing a plurality of phonetic sub-units in *Chou et al.*

Therefore, the rejections of claims 1 to 5, 21 to 25, 30 to 32, 37 to 39, 48 to 52, 68 to 73, 77 to 79, 84 to 86, 90, 97, 99, and 101 under 35 U.S.C. 102(b) as being anticipated by *Chou et al.*, of claims 20, 35, 36, 46, 67, 82, 83, 93, 98, 100, and 102 under 35 U.S.C. 103(a) as being unpatentable over *Chou et al.* in view of *Aref et al.*, and of claims 33, 34, 80, and 81 under 35 U.S.C. 103(a) as being unpatentable over *Chou et al.* in view of *Wheatley et al.*, are proper.

Conclusion

8. **THIS ACTION IS MADE FINAL.** Applicants are reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Martin Lerner whose telephone number is (703) 308-

9064. The examiner can normally be reached on 8:30 AM to 6:00 PM Monday to Thursday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (703) 305-9645. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

ML
3/22/05

Martin Lerner
Examiner
Art Unit 2654